

GENETIC PROFILING AND BANKING SYSTEM AND METHOD

Background of the Invention

This invention relates to a system for profiling and banking genetic information about individuals.

5 Genetic information derived from a sample can be used to determine medical and other information about an individual. Obtaining genetic samples and testing those samples raises concerns about privacy, access, and the use of the resulting information. Such information could, however, be useful for individuals and medical practitioners.

10 Currently there are companies that provide banking services for storing physical samples of genetic material from biological tissue, such as blood or cheek cells. That material can be retrieved from storage and tested as desired. Removing the sample and having it tested is time consuming and may be repetitive if multiple tests are needed over time.

15

Summary of the Invention

The present invention includes systems and methods for storing and accessing genetic information. The systems and methods preferably provide protection against unauthorized access and use, but provide convenience in
20 accessing and using genetic information if such use is properly authorized.

In a method according to the present invention, a patient enrolls in a genetic banking system and provides a source of genetic material, such as a blood sample. The sample is processed using a combination of experimental and/or *in silico* techniques to produce a genetic profile for the patients. The processed data is
25 stored in a database to create a genetic profile for that patient. A remaining portion of the physical sample may also be stored for further use if needed.

The patient, either at the time of enrollment or after a sample is processed, can dictate access rights, including the ability of third parties (other than the

individual or the genetic bank itself), such as medical practitioners, to access this profile, and the specific purposes for which the profile can be accessed and used. Thus, the patient can specify both the people who will have access, and the uses for which they have access.

The system provides user interfaces for the user to enter identification
5 information and access rights. The bank can prepare standard protocols that describe allowed and proscribed data sharing.

A medical practitioner authorized to have access by the patient and confirmed by the access control system can access the profile that is stored and can run tests based on the profile as stored in the database; for example, such a test can
10 be run to determine the presence or absence of certain markers.

In another aspect, a system according to the present invention includes a database for storing genetic data on individuals and an access control system that controls access to the database and manages the tests that are to be done. The control system interacts with (or includes) a testing system to cause the testing
15 system to process the profile data to conduct the desired test. The control system authorizes the test and provides the results.

The system allows users to store a comprehensive digitized DNA profile based on a sample, in addition to storage of physical samples. The patient has control to voluntarily allow access to particular people and for particular purposes,
20 thus protecting the privacy of that information. Because the samples have been processed and digitized, additional tests can be performed without requiring repetitive use of actual physical samples. Other features and advantages will become apparent from the following description of preferred embodiments, drawings, and claims.

25

Brief Description of the Drawings

Fig. 1 is a block diagram of a system according to the present invention.

Figs. 2-4 are flow charts showing methods according to the present invention.

Description of Preferred Embodiments

A genetic banking system 10 has a database 18 for storing genetic profiles, and a processing system 12 for controlling interactions between a patient 14 and database 18, and between a medical practitioner 16 and database 18. System 10 may
5 also include data integration and analysis functionalities and a testing system 20 that is linked to database 18, and may also include physical storage 22 for storing physical samples of genetic material. Processing system 12 maintains access rights that allow a patient to specify and control the access to the profile by others such as medical practitioners, enforces the access rights that are voluntarily provided from
10 the patient, causes testing system 20 to perform a desired test, and provides results to the medical practitioner.

Database 18 may be a relational database with a relational database management system (RDBMS), or it may use technology such as that in ACEDB, a genome database system that has a custom database kernel, a data model designed
15 specifically for handling scientific data flexibly, and a graphical user interface with many specific displays and tools for genomic data. ACEDB software and supporting tools are publicly available via the Internet by download.

Processing system 12 and the testing system 20 may each include a special purpose computer, a workstation, a server, or some combination of linked
20 computers, workstations, and/or servers for interfacing with users, processing, information and performing tests. The connections employed are preferably high speed, TCP/IP connections, such as T1 lines. The testing system may be independent and remote from processing system 12 that interacts with users and medical practitioners, or the testing system and processing system may be part of
25 one larger processing system.

Referring to the flow chart of Fig. 2, a source of genetic material, such as a blood sample, is provided by a patient for testing purposes after the patient enrolls with the system (100). This enrollment process includes the processing system

providing a user interface to the patient, and the patient providing information about himself/herself through the user interface to the processing system about himself/herself. The patient, using the user interface, voluntarily establishes conditions under which the genetic information can be accessed, preferably through use of a menu of standard protocols or by selecting a standard protocol and customizing it. The patient can thus determine access rights, including both who gets access and the uses that can be made of that genetic information. Thus, at least some authorized third parties users may have access for some purposes and not for other purposes, or some third parties may have access rights set to "all" for all purposes.

10 To confirm selections regarding access and use made by the patients, the system preferably employs a process of querying and confirming through the user interface, and preferably also includes information for the patient about the system and about the tests. Redundancies and checks can be provided in the interaction between the patient and the user interface during the enrollment process so that the patient understands the possible uses of the profile and the consequences of selecting who gets access and the uses for that access.

To enroll, the patient preferably either obtains from the processing system a password at the time of enrollment, or the patient may have previously been provided with a password for confidentiality. The system preferably provides a high level of security and may include mechanisms such as digital certificates in addition to the password protection. The enrollment can take place on a private or other closed or dedicated network, such as a LAN or Intranet, or with appropriate security measures, over the Internet, and can be performed without additional human intervention (i.e., additional to the user). For Internet interaction, the processing system can include a web server for providing pages or forms and for receiving information entered onto those forms.

When the patient is enrolled, the patient has a physical source of genetic material shipped to the genetic banking system or some other desired location for processing (102).

At a later time, the patient can use the password or a new password to change selections for who can get access and the uses for which access can be made. In addition, the selections can "time out" such that the user must re-confirm selections after some period of time or else the access rights are terminated.

The physical sample is processed by the testing system and the results of the processing are stored in the database. The processing that is performed on the sample can vary depending on the genetic banking and testing services that are provided, but can include genotyping and bioinformatic profiling of general and/or specific genetic marker panels. Such information can be used to determine risks of many diseases including, without limitation, cancer, Huntington's Disease, Alzheimer's Disease, and hypertension. The tests can be done through a number of different methods, such as fluorescence, optical density, mass spectroscopy, DNA sequencing, microarray-based assays, or other current methods or methods to be developed. The data that is provided from these tests is digitally stored in the database as a genetic profile of the patient for subsequent analysis and tests.

After the tests are done, the actual physical sample (or at least any remaining portion) may be stored if desired in case it is needed for confirmation or other future purposes (106). This physical storage can be done using known cryogenic techniques.

Referring to Fig. 3, a third party, such as a medical practitioner (i.e., a party other than the patient or people associated with the management and/or operation of the genetic banking system), may seek access to an enrolled individual's profile in the database (110) using a user interface provided from the processing system to the practitioner. The practitioner is authenticated, such as through the use of digital certificates and/or password protection before he/she can access the information. The system compares the access rights entered by the patient into the system with

that third party practitioner seeking access to the patient's profile. The system then determines whether access is allowed for that third party practitioner (112).

If the access is allowed, the practitioner can seek to have a test performed, such as a search of the profile for markers for Huntington's Disease (114). The system determines whether the test is authorized by comparing the type of test that is desired with the access rights entered by the patient (116) (a practitioner may have access for some purposes but not others). If the test is authorized, the processing system forwards the request to the testing system to perform the test. The results of the test can be analyzed in the processing system if needed and provided to the practitioner, preferably in an encrypted form (118). The connections that are provided, particularly between database 18 and testing system 20, should have high bandwidth to allow a significant amount of data to be quickly provided to the testing system for the results. Testing system 20 should be sufficiently computationally intensive in order to perform the requested tests.

If access by the practitioner is not allowed, or if the particular test is not authorized, the patient can log in to the system to provide the appropriate authorizations for access (120, 122).

The practitioner can seek access over a dedicated or closed network, or with appropriate security, over the Internet. The processes of seeking access, confirming access, performing the test, and reporting results can all be performed without additional human intervention (other than the medical practitioner himself/herself).

The patient may or may not have access to have tests run individually and without the tests being performed for a medical practitioner. It may be desirable for the patient only to have access through a medical practitioner who can explain to the patient the actual meaning of the tests. However, it would also be possible to allow the patient to have tests performed and to have access to the results.

Referring to Fig. 4, a more detailed example of the sample processing process is shown. In this process, the DNA is extracted from the sample, such as the blood sample (130). Extraction of DNA from a sample of cells may be accomplished by

(104)

any standard method. Using the patient's DNA sample, a complete genotype (GT) and pan-loci ID are performed (132). For example, a single nucleotide polymorphism (SNP) map may be generated from the patient's DNA sample. In the alternate, if the patient has or is predisposed to develop a specific disease, genes (e.g., mutations, aberrant expression patterns) associated with those diseases may be directly sequenced from the patient's sample. Next, high-throughput (HT) genotyping is executed (134). The results of the genotyping data are provided in reports that are preferably customized for convenient use (136).

Several non-limiting genes, and the diseases with which they are associated, which can be sequenced and/or from which a SNP map can be generated according to the methods of the invention are provided in Pulst S.M. (1999) *Neurol.* 56(6): 667-672; Ballantyne et al. (1997) *J. Cardiovasc. Risk* 4(5-6):353-356; Marian, A.J. (1997) *J. Cardiovasc. Risk* 4(5-6): 341-345; Marian, A.J. (1997) *J. Cardiovasc. Risk* 4(5-6): 333-339; Hallman, D.M. et al. (1997) *J. Cardiovasc. Risk* 4(5-6): 325-331; Ballantyne et al. (1997) *J. Cardiovasc. Risk* 4(5-6): 321-323. Additional genes the diseases with which they are associated are listed in Table I.

Table I

Gene	Disease	Reference*
p53	cervical cancer	Zehbe et al. (1999) <i>Lancet</i> 354(9174): 218-219
BRCA1	breast cancer	Hakansson et al. (1997) <i>Am. J. Hum. Genet.</i> 60(5): 1068-1078
BRCA2	breast cancer	Hakansson et al. (1997) <i>Am. J. Hum. Genet.</i> 60(5): 1068-1078
CTLA-4	thyroid associated orbitopathy	Vaidya et al. (1999) <i>Lancet</i> 354(9180): 743-744
nitric oxide synthase (eNOS)	coronary artery disease	Liao et al. (1999) <i>Angiology</i> 50(8): 671-676
gene encoding human ATP cassette-binding	Tangier disease	Rust et al. (1999) <i>Nat. Genet.</i> 22(4): 352-355; Brooks-Wilson (1999) <i>Nat. Genet.</i> 22(4):336-345

transporter 1 (ABC1)		
CTLA-4	multiple sclerosis	Ligers et al. (1999) J. Neuroimmunol. 97(1-2): 182-190
genetic marker D6S461	multiple sclerosis	Shinar et al. (1998) J. Mol. Neurosci. 11(3): 265-269
Microdeletions at chromosome bands 1q32-q41	Van der Woude syndrome	Schutte et al. (1999) Am. J. Med. Genet. 84(2):145-150
chromosome 21q22	bipolar affective disorder	Aita et al. (1999) Am. J. Hum. Genet. 64(1):210-217.

*each reference is hereby incorporated by reference

10

15

20

In addition to specific genes associated with specific diseases, many diseases are characterized by their association with a set of at least one genetic marker which can be detected using the methods of the present invention. For example, a marker in the UCP-2/UCP-3 gene cluster has been linked to a genetic susceptibility to anorexia nervosa (Campbell et al. (1999) Mol. Psychiatry 4(1): 68-70). Likewise, genetic markers in addition to apolipoprotein E (APOE) polymorphism has been associated with Alzheimer's disease (Scacchi et al. (1999) Neurosci. Lett. 259(1): 33-36). Parkinson's disease is similarly associated with a certain combined alpha-synuclein/apolipoprotein E genotype (Kruger et al. (1999) Ann. Neurol. 45(5): 611-617). The methods of the invention may also be employed to detect the presence of a multigenic disease (or the predisposition to develop such a disease) that is associated with and/or caused by the allelic variants of more than one gene.

25

For the purposes of the invention, it matters not whether the disease is caused by and/or correlated with the associated genetic marker or whether the associated genetic marker is caused by and/or correlated with the disease. What is relevant is that certain genetic markers, such as allelic variants, are associated with certain disease phenotypes and/or predisposition to develop the disease.

The invention provides an accessible confidential database that creates and stores genetic information, such as an SNP map, from patient DNA. As more patient

samples are added to the database of the invention, and as researchers find more associations between genetic markers and particular disease phenotypes, use of the invention provides an on-going self-perpetuating advancement into the development of associations between genetic markers and certain diseases.

Moreover, with appropriate security and privacy cautions, the database can be used
5 in an anonymous manner to allow researchers to access a body of genetic information for research and analysis purposes.

The system of the present invention thus allows a patient to voluntarily bank genetic information that can be used quickly to determine genetic and medical information about that individual, particularly information that relates to whether
10 the individual carries genetic information associated with known diseases. The system provides restrictions, however, that allow the user to retain privacy and limit unauthorized access to his/her genetic information. The system is thus unlike a system, for example, in which DNA information, like fingerprint information, is stored for identification purposes to use DNA information to identify individuals
15 involved in specific criminal activities; in such a case, the individual who provides the sample would generally not have voluntary control to establish the ability of others to access the information, and such systems would generally not have the ability to test for a number of different medical purposes for which the DNA information can be accessed by others.

20 The patents and scientific literature referred to herein establish the knowledge of those with skill in the art and are hereby incorporated by reference in their entirety to the same extent as if each was specifically and individually indicated to be incorporated by reference. Any conflict between any reference cited herein and the specific teachings of this specification shall be resolved in favor of the
25 latter. Likewise, any conflict between an art-understood definition of a word or phrase and a definition of the word or phrase as specifically taught in this specification shall be resolved in favor of the latter.

Having described a preferred embodiment of the present invention, it should be apparent that modifications can be made without departing from the scope of the invention as defined by the appended claims. For example, the authorization is described as a two-step process of checking the user then the purpose, although the authorization could combine user and purpose into one combined access right.

- 5 While components of the preferred embodiment may have certain benefits and advantages noted herein, other systems and components thereof may nonetheless be within the scope of the claims without necessarily having each and every one of the noted benefits and advantages.

What is claimed is:

10

15

20

25